

## Science ethics as a bureaucratic problem: IRBs, Rules, and Failures of control\*

BARRY BOZEMAN & PAUL HIRSCH

*Research Value Mapping Program, School of Public Policy, Georgia Tech, Atlanta, Georgia, U.S.A.;*  
*E-mail: barry.bozeman@pubpolicy.gatech.edu*

**Abstract.** “Institutionalized science ethics” refers to the statutory, professional and institution-based ethical standards that guide and constrain scientists’ research work. The primary institution responsible for implementing institutionalized science ethics is the Institutional Review Board. We examine the limitations of IRBs and institutionalized science ethics, using bureaucratic theory and, especially, theory related to the development and enactment of rules. We suggest that due to the very character of rules-based systems, improvements in IRB outcomes are unlikely to be achieved through either more or better rules or even by bureaucratic reform. Instead, we suggest that improvements in human subject protection can best be advanced through increased participation. Ours is not a call for more participation by the general public but participation, via “Participant Review Boards” of persons who are eligible, by the protocols of the research in question, to serve as subjects. This provides a level of legitimacy and face validity that cannot be obtained by IRB affiliates, even by “external representatives.” In making these points, we review a recent science ethics controversy, the KKI/Johns Hopkins lead paint study. In spite of being approved by IRBs, the study resulted in a civil lawsuit that reached the Maryland Court of Appeals. The case illustrates the limits of institutionalized science ethics and the bureaucracies created for their enactment. The case also underscores the complex and equivocal nature of the ethical guidelines established under the National Research Act.

*“Thirty years ago, one out of every three children admitted to KKI’s hospital was admitted for lead poisoning, and every year for decades a dozen children died of it. Because of our research and pursuit of achievable solutions to this epidemic, no children have died of lead poisoning in the last several decades in Baltimore... [At the time of the 1992 study] children were still being poisoned with alarming frequency... We knew how to stop the poisoning, we knew what to do, but we needed to conduct the 1992 study to demonstrate how well these prevention methods worked. And that’s what the study accomplished.”*

*-Gary Goldstein, Director of KKI*

*The researchers and their Institutional Review Board apparently saw nothing wrong with the research protocols that anticipated possible accumulation of lead in the blood of otherwise healthy children as a result of the experiment.*

*-Maryland Court of Appeals*

\*We are grateful to – for providing helpful comments on an earlier draft.

## Introduction

*Baltimore, Maryland 1992.* Researchers from the Kennedy Krieger Institute (KKI), an affiliate of Baltimore's Johns Hopkins University, propose a study to evaluate the effectiveness and longevity of various measures for reducing the amount of lead that children are exposed to in their homes (Kaiser, 2001; Lewin, 2001). Despite a 1978 Federal ban on the use of lead in household paint, paint chips and lead paint dust present in homes built before 1978 pose a significant threat to children. In inner-city Baltimore in 1992, up to 95% of the homes in some of the poorest neighborhoods contain lead.<sup>1</sup>

In the decades leading up to the study, researchers at the Kennedy Krieger Institute have played a leading role not only in treating childhood lead poisoning, but also in bringing the dangers of childhood lead paint exposure to the public consciousness. Additionally, KKI researchers have been instrumental in critiquing half-hazard yet generally accepted techniques for removing lead paint, and in developing and testing effective alternative techniques for lead paint abatement.<sup>2</sup> The 1992 study, a federally funded outgrowth of this work, employs a research methodology that involves applying several lead abatement techniques to rental homes in Baltimore's high risk areas, and monitoring the blood lead levels of the children living there.

The KKI study proceeds, but only after the study's methodology and procedures have been reviewed and approved (Adelman et al., 2001) by the U.S. Environmental Protection Agency and by the Johns Hopkins University Institutional Review Board (discussed below), the committee endowed with the authority to rule on the ethics and acceptability of research involving human subjects. Throughout the course of the research, the majority of subjects experience reductions in blood lead levels (Lewin, 2001).

*August 21, 2001.* Following months of controversy and acrimony, the participants and researchers in the KKI lead paint study open the *Washington Post* to read the headline "Maryland Appeals Court Slams Researchers; Participants in Study on Lead Paint Weren't Informed of Risks, Judge Says" and a lead paragraph worthy of a tabloid's purple prose:

*The Maryland Court of Appeals' sweeping condemnation of lead paint researchers at Baltimore's Kennedy Krieger Institute has reinforced the rights of research subjects to know the risks they face, while tainting the prestigious research center by comparing its work to the infamous Tuskegee, Ala., experiments [our emphasis] that withheld treatment to black men infected by syphilis (Roig-Franzia and Rick Weiss, 2001, p. B01).*

The suit (Grimes vs. KKI) reported in the *Post* account that brought the lead paint studies to a grinding halt, and that shone a national media spot light on the researchers and institutions involved, was initiated by two mothers whose sons' blood lead levels became elevated during their participation in the study (Kaiser, 2001). Public outrage followed, fueled by the revelation that in some of the houses in which experimental abatement measures were applied, KKI researchers had encouraged the landlords to rent to families with young children (Kaiser, 2001).

## **Institutionalized science ethics**

Was the KKI study “unethical?” If so, are the results an indictment of the institutions that have been established for adjudicating science ethics? Were unfortunate outcomes, whether or not ethical, potentially avoidable with more prudence, more safeguards or better bureaucratic procedures? Overall, the case is excellent grist for considering contemporary institutional arrangements for determining science ethics.

“Institutionalized science ethics” refers to the statutory, professional and institution-based ethical standards that guide and constrain scientists’ research work. We distinguish these standards from, on the one hand, the personal ethics of individual scientists and, on the other hand, broader statutory constraints and legal standards.<sup>3</sup> Nor do community standards, public opinion, or public expectations qualify as institutionalized science ethics. Institutionalized science ethics:

1. are formal and codified;
2. address specifically possible harm and benefits from research;
3. are broadly communicated within the research community and widely viewed as legitimate by that community;
4. include some compliance or enforcement mechanisms.

Today, the most recognizable institution for codifying and implementing institutional ethics is the Institutional Review Board (IRB). The IRB is the mandatory bureaucratic mechanism for reviewing all proposals submitted for scientific research involving human subjects. The IRB is the first institutional reviewer of science ethics but, of course, most researchers, in designing studies that involve humans as subjects, have already weighed the hoped for benefits of the research against its potential risks; they also may have given careful thought to how to obtain informed consent from research subjects; they may even have given consideration to matters of justice and equity – each principles that IRBs are designed to uphold. Certainly few, if any, researchers actively intend to exploit or cause harm to human subjects participating in their research. Nevertheless, history has shown that a researcher’s forethought or good intentions are not enough to prevent exploitative research; thus, since the mid 1970s Institutional Review has been required and regarded as the primary safeguard in preventing exploitation or harm to human subjects.

As a form of institutionalized ethics, IRBs have many advantages, but they also – at least as currently construed – have limitations. They may be particularly limited in evaluating the merits and risks of research that draws participants from vulnerable populations – minorities, the poor, the handicapped, prisoners, children, etc. – who are extremely unlikely to be adequately represented by IRB committees. Furthermore, they do not generally consider in depth the scientist’s effectiveness in communicating with potential subjects so that, if given, consent is in fact “informed.” In this paper, we make the case that although Institutional Review Boards are an essential line of defense against abusive or exploitative science, complementary means for protecting research subjects may be warranted for certain kinds of research that involves vulnerable populations. Our objectives in this paper are to analyze the limitations of IRBs from

an institutional perspective, and to propose and discuss a complementary safeguard – *Participant Review Boards* – for vetting certain kinds of human subjects research.

The paper will develop as follows. First, we define and briefly consider the development and evolution of institutionalized science ethics over the last several decades, with particular attention paid to the adoption, structure and function of IRBs. Next, we illustrate the limitations of IRBs by presenting key details of a specific research study that was labeled unethical by a State Appeals Court *in spite of* careful vetting of research plans and procedures by an IRB. Third, we draw on recent scholarship regarding rules and their application to develop a theoretical understanding of the limitations of IRBs. Finally, we propose and articulate a role for Participant Review Boards as a complement to current approaches to institutionalized science ethics.

### **A brief history of institutionalized science ethics**

To understand the current state of affairs and dominant institutions for science ethics it is helpful briefly to consider the history of institutionalized science ethics. It is a history characterized by disaster response. The content of the drama changes, but it is always a three-act play. Act One: scientists, either through malevolence, ignorance or misguided good intentions engage in (allegedly) exploitative or harmful research; Act Two: these apparently unethical practices come to light and are roundly criticized; Act Three: policies, guidelines, or ethical codes are developed to reduce the likelihood of reoccurrence.

While moral codes for the practice of research date back to the Hippocratic Oath, modern institutionalized science ethics can be traced to the aftermath of World War II. In 1946, the Nuremberg trials focused attention on human experimentation in the prosecution of twenty-three German scientists, physicians and administrators for their culpability in infecting or performing unneeded surgery on thousands of persons incarcerated in concentration camps. These human experimentation trials gave rise in 1948 to the Nuremberg Code, which was one of the first codified statements mandating what has come to be called “informed consent.” The Code also articulated the (still murky) principal that the benefits of research must outweigh the risks. While the Nuremberg Code itself does not have the force of law, many elements of current human subjects’ norms, ethical codes, and laws derive from the Nuremberg Code.

In one of the few instances of an ethical declaration or policy not precipitated by some identifiable human tragedy, in 1964 the World Medical Association<sup>4</sup> promulgated the set of human subjects standards commonly referred to as the Helsinki Declaration (Enserink, 2000). While the Declaration is in many respects similar to the Nuremberg Code, it is much more detailed. Compared to the Nuremberg Code, the impact of the Declaration of Helsinki seems to be much more extensive, especially in terms of practical application (Howard-Jones, 1982; Human and Fluss, 2001). Until the passage of the National Research Act, the Helsinki Declaration was the world’s most influential set of institutionalized science ethics.

Perhaps the most important developments in the history of institutionalized science ethics emerged in the wake of the political uproar from public airing of the Tuskegee studies<sup>5</sup>, which were begun in 1932 and finally put to a stop in 1972 after newspaper

articles were published documenting that African American sharecroppers diagnosed with syphilis were purposely left untreated despite the known effectiveness of penicillin. In 1974, the National Research Act was passed to prevent abuses such as those promulgated in Tuskegee from ever happening again. The National Research Act was the basis for the creation of the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which was charged with developing guidelines for protection of human subjects. The Commission's *Belmont Report* is one of the landmark documents of institutionalized science ethics. The central component of the National Research Act is to require that all research protocols involving human subjects be reviewed and approved by an Institutional Review Board (IRB).

Three core ethical guidelines were the centerpiece of the *Belmont Report*. The principle "*respect for persons*" holds that (1) all individuals should be treated as autonomous agents, and (2) persons with diminished autonomy are entitled to protection. The "*beneficence*" principle states that persons should be treated ethically not only by providing autonomy and protection from harm but also that efforts should be made to ensure their well-being. This standard goes further than the Hippocratic injunction to "do no harm" by calling for research to positively impart benefits to participants. The beneficence principle recognizes that some harm may be a necessary pre-condition to benefit and, thus, the beneficence principle endorses a risk-benefit calculation that is at the center of some science ethics controversies. The third principle, "*justice*," considers who receives the benefits of the research relative to who bears the burden.

The Belmont principles are operationalized in the Code of Federal Regulations, Title 45 CFR Part 46, Protection of Human Subjects, also known as the Common Rule. The Common Rule obtains for all scientific research funded by the U.S. federal government and has been adopted by all federal agencies conducting or funding human subject research (thus the *Common Rule*). While the regulations pertaining to the common rule are voluminous (see NIH, 2005) they are based on three requirements: (1) compliance by research institutions, (2) assurance that researchers will obtain and document informed consent from all research participants, and (3) the establishment of IRB obligations including membership, function, operations, review of research, and record keeping.

#### *The IRB: Science ethics bureaucratized*

In the United States, more than 6,000 IRBs are dispersed throughout institutions that conduct federally funded research on human subjects (Gray, 1978; Lane, 2005). IRBs are attached to universities and federal agencies, but they are also found in hospitals, foundations, and private research organizations.<sup>6</sup>

The duties of IRBs vary little, and all pertain to the protection of human subjects. However, regulations regarding IRBs permit considerable latitude with respect to the membership and organizational structure of IRBs. For example, the Food and Drug Administration, one of the agencies overseeing an especially large number of IRBs, stipulates only that "an IRB is an appropriately constituted group that has been formally designated to review and monitor...research involving human subjects" (FDA, 2005). The IRB does not need to be attached to the institution that is doing the research under review, and some research institutions rely entirely on IRBs affiliated

with other institutions. Limitations on memberships are minimal, chiefly prohibiting conflict of interest. Otherwise, the only requirement is that at least one member not be affiliated with the institution housing the IRB, at least one member must be a scientist or clinical researcher, and at least one member must be a non-scientist. In some cases one individual can satisfy two categories – for example, being a non-scientist and lacking affiliation with the institution (to the extent that outside members retain their IRB seat for prolonged periods they may come to view themselves as being to some degree affiliated with the institution).<sup>7</sup> In addition to these formal requirements, IRBs are encouraged, absent specific requirements, to promote diversity, “including consideration of race, gender, cultural backgrounds and sensitivity to such issues as community attitudes” (FDA, 2005).

Relatively little systematic research on IRBs has been published. Most research is either based on IRB members’ perceptions (e.g. Cleary, 1987; Ferraro et al., 1999) or focuses on descriptive aspects of the IRB including such factors as periodicity of meetings, workload management, and outcomes of reviews (e.g. Gray et al., 1978; Goldman and Katz, 1982; Bell et al., 1998). Barber and colleagues (1973) published what is perhaps the first major study of IRBs, research using questionnaires returned from respondents from more than 300 biomedical IRBs and reporting on their review outcomes. A primary focus was on the degree of consistency among IRBs with respect to a willingness to sustain risk. The researchers presented several hypothetical risks to respondents and found that a relatively small number of respondents were willing to sustain a high level of risk. However, the reliance of the study on scenarios presents some possibility of socially desirable response bias.

In one of the best known IRB studies, Gray and colleagues (1978) examined 61 IRBs in a study focusing on performance and outcomes. Using questionnaires, the researchers examined the work details of IRBs including frequency of meetings, attendance, decisions about submitted protocols and members’ perceptions of the extent to which they were qualified to pass judgment on the studies they examined.

Most researchers focusing on IRBs treat them as unique institutions, which, of course, they are in some respects. At the same time, IRBs share many characteristics with other entities, including a more or less stable and identifiable structure, explicit goals, authority lines, and, especially important for present purposes, governance by formalized rules, regulations and procedures. As an institution that combines structural formality and reliance on rules with the flexibility that comes from decentralization, the IRB has some important strengths including: (1) some of the participants are drawn from the broad research community and are themselves researchers; (2) there is at least some attention to including persons who are not part of the research institutions; and (3) while not universally loved, the IRB is universally accepted.

All institutional arrangements have their limitations, however. Perhaps the most important limitation in the case of IRBs is the limited involvement of persons who directly represent research subjects or who are themselves potential research subjects. While IRBs must have members drawn from outside the institution they are a part of, the “outside member” is often a professional such as a minister, teacher, or social worker.<sup>8</sup> There is no guarantee or even expectation that the “outside members” do or could represent research subjects, particularly when those research subjects are members of vulnerable populations.

In the next section, we present the details of a case study in which an IRB application was submitted by researchers and approved, and which after being implemented resulted in a lawsuit and allegations of abuse and exploitation, including direct comparisons by five State judges to the infamous Tuskegee studies. While many interpretations of the case are possible, we use it to illustrate the limitations of IRBs that accrue from their institutional, rule-based structure and their lack of empathy and/or ability to communicate with the – in this case – poor and minority populations who participated in the research and gave their consent based on the assumption they had been given sufficient information to be informed.

### **The KKI case: Distilling lessons about science ethics and political controversy**

*“Kennedy Krieger cannot understand how a study funded by the federal government and in compliance with federal regulations can be said to be unethical in its design.”*

*–Attorneys for the Kennedy Krieger Institute, an affiliate of Johns Hopkins University, in a legal brief filed by the defense in a suit against University researchers (Adelman et al., 2001).*

Lead poisoning is often considered the most common environmental disease of young children (Agency for Toxic Substances and Disease Registry, 1988), and research has shown that about one in five children exhibit dangerous levels of lead, and nearly 900,000 have sufficient concentrations of lead to pose a serious threat of mental retardation and other brain disorders (Centers for Disease Control, 2004b). Currently, the major source of lead poisoning for children is the deteriorating lead paint that was used in houses built before 1978, when lead paint was banned. The Centers for Disease Control (CDC) estimate that 24 million houses with lead paint still exist in the United States, at least 4 million of which are inhabited by young children (CDC, 2004b). Not surprisingly, the families living in such houses are predominantly lower-income, members of minority groups, or both (National Research Council, 1993; Lamphear, Weitzman et al., 1996; Bernard, 2004).

The problem is well known and its treatment a high priority. In 1991, the U.S. Department of Health and Human Services called for a society-wide effort to eliminate childhood lead poisoning in 20 years. The KKI study emerged as part of that strategy. The KKI project focused on the cost of lead paint remediation. While cost is certainly not the only factor involved in reducing children’s exposure to lead paint, there seems little question that it is an important one.<sup>9</sup> It is important to emphasize, as mentioned above, that the Johns Hopkins University IRB reviewed and approved the research. Under the research procedures, the participants, all residents of low income homes in Baltimore, received one of three abatement measures, and the blood lead levels of children living in those houses were intermittently measured (Lewin, 2001). Following widely accepted ethical guidelines the KKI researchers ensured that all of the houses in the study received some form of lead abatement (i.e. there was no placebo).

The majority of the children experienced reductions in blood lead levels (Lewin, 2001). But not all of them: The law suit which ultimately led to the Maryland Court of Appeals condemnation of the KKI study was motivated in part by the fact that some children had higher levels of lead than when the study began. However, the most poisonous publicity occurred in the wake of revelations that the KKI researchers had encouraged the landlords of some of the lead abated houses to rent to families with young children (Kaiser, 2001). While Johns Hopkins maintains that all of the families recruited for the study had a history of living in lead contaminated housing<sup>10</sup>, this justification wasn't convincing enough to prevent the Appeals Court from accusing researchers of using kids as "canaries in the coal mine" (Curry, 2001).

The KKI lead paint script can be read from many different perspectives (Ross, 2002; Spriggs, 2004), including (1) the cultural divide between apparently well-meaning elite scientists and the disadvantaged public they wish to serve, (2) the difficulties of enacting ethical judgments in the political arena, or (3) problems in public understanding of science. However, our focus is on the implications of the lead paint case for what we call "institutionalized science ethics" and, particularly, the limitations of bureaucratic control exercised through rules and procedures. A key point to our study is that the KKI lead paint controversy occurred despite the many safeguards and controls that have been established to safeguard human subjects. An Institutional Review Board vetted the KKI project's study procedures and the study was judged acceptable within the canons of ethics and under the rules developed for IRBs. While opinions differ as to the ethical implications of the KKI study's research, there seems little doubt that there is a disjunction between institutionalized science ethics and public expectations: that an "appropriate and properly designed" study would either involve no risk to participants or would communicate any known risk to those participants ahead of time.

Regardless of one's position on the ethics of the case, the following points seem clear: (1) The researchers and the Johns Hopkins IRB arrived at very different conclusions than the Maryland Court of Appeals and the plaintiffs in the lawsuit, and, (2) whether or not the KKI case is a "new Tuskegee," it is noteworthy in its contribution, with the complicity of a media more interested in sensationalism than balanced analysis, to the ever widening gap between scientists and those who are either poor, less educated, minority or in some other respect vulnerable. This is particularly salient given the observation that the KKI researchers were specifically intending to *help* vulnerable individuals. A Johns Hopkins researcher knowledgeable with the study told us that, in the opinion of the researcher, the court's comparison with Tuskegee was particularly unhelpful not only because it was sensationalistic, but also because it made the researchers' scientific colleagues (arguably the intended audience of the court's decision) basically dismiss the court's opinion as ridiculous and therefore not worthy of consideration.<sup>11</sup>

For the sake both of researchers that choose to work on ways to ameliorate situations where significant inequalities occurred *before* scientists ever got on the scene, as well as for the sake of vulnerable populations that live in situations characterized by inequality, it is important to sort out what went wrong with the accepted institutional process for preventing just the type of lose-lose scenario that unfolded.

*What, if anything, is distinctive about the kki case?*

The KKI case is not at all clear-cut and, in all likelihood, there are several hundred studies being conducted right now that have ethical dilemmas just as complex as the KKI case but, as yet, have not had the “misfortune” to have to confront activist subjects, legal attendants, and an alarmed jurist. In our judgment, KKI resembles a great many other cases in the following ways:

1. In KKI, as in a great many studies, there was a trade-off between, on the one hand, significant risk and knowable harm to subjects and, on the other, benefits to both the subjects and, potentially, to society.
2. As is the case in most studies vetted by IRBs, there was a clear intention by KKI researchers to provide information about their study, its protocol, and its possible consequences, yet the information provided may not have been communicated completely or effectively. The usual standard for informed consent, from the *Belmont Report*, requires provision of information, comprehension and volunteering (Dunn and Chadwick, 1999). None of these is easily adjudicated: few subjects, even quite sophisticated subjects, comprehend the full intent, purpose and practices of any research project and, moreover, the rules permit some degree of deception. Nor is the act of volunteering always clear-cut. It is easy enough to determine if someone has signed a document, but social psychology has shown us the many different ways in which individuals can be influenced by the way in which a choice is communicated. Rarely – if ever – do IRBs question and/or evaluate the scientist’s ability to engage in such communication with their subjects.
3. KKI researchers provided some direct therapeutic treatment to many (but not all) of the research participants.
4. Fourth, researchers addressed a social problem of some considerable importance with a research objective that had some probability of effectively addressing the problem.

There are two important ways in which the KKI case differs a great deal from the most studies. In the first place, the KKI case addresses head on an issue that is of particular interest to minorities, pregnant women, low-income families, and children-vulnerable subjects all. The Johns Hopkins study sought to address needs of the poor (middle class children also suffer from lead poisoning, but at much lower rates). Social critics (Farmer, 1999; Campbell, 2000) and public policy-makers (Brown, 1993; Johnson et al., 1998) decry the paucity of research addressing the needs of the poor. The Johns Hopkins lead paint study seems to be exactly the sort of research that is in line with government, especially NIH, objectives focusing research on the needs of underserved, poor and/or minority populations. We allege that scientists who do study these vulnerable populations generally face an increased risk of being judged “unethical.” Thus, one of the crucial reasons to avoid the harmful fallout of the sort that occurred in KKI is to discourage researchers from always taking the safe route: working on problems not especially of interest to vulnerable populations.

A second crucial respect in which the KKI study was different from the thousands vetted daily by IRBs is in its active recruitment of research subjects who are already embedded in a situation that is highly charged politically, and that is characterized by pre-existing dramatic inequalities. Furthermore, it is researchers from that same institution, many of the exact same researchers involved in the condemned study, who over the past two decades have shed light on the dangers of lead paint for the development of children, making them in effect champions of those most at risk of lead exposure.

But it is one thing to shed light onto a social problem and another to try to fix it, and the scientific method may be more attuned to the former rather than the latter objective. Scientifically, it makes sense that the real-life success of lead paint remediation can only be assessed according to how it limits exposure in actual circumstances. Furthermore, given the preponderance of lead paint in the low-income Baltimore housing stock, it is easy to imagine the logic of a researcher or an IRB member who assumes that since all of the children in the study would *probably* (key word) be living in houses with lead paint with or without the research, and since the research provides the added benefit of abatement and monitoring, it is OK to encourage landlords to rent experimentally abated homes to families with young children.

But this is the ethical rub. The proposition that the recruitment technique employed is justified because children would have likely been exposed to lead paint anyways is the one assumption easily impeached by examining even the most general of ideals embodied by the original ethical mandate for human subjects, the Hippocratic Oath—“First, do no harm.” “First do no harm” does not say or mean: “Do no harm that would not otherwise be done in the natural course.” It is an injunction to refrain from doing harm, not an invitation to do benefit-cost analysis or risk analysis. As improbable as it may seem, some of the people who signed their children up for the research may have, if adequately informed of the risks, been able to find an alternative dwelling, or perhaps they might even have been moved to political activism on behalf of the entire community. Therefore, if we factor in the problems of effectively communicating risk to vulnerable populations (not to mention the difficulty of communicating to members of disadvantaged groups that it is truly OK *not* to participate), the Maryland Court of Appeals condemnation of the study is as understandable as the oversight by the Johns Hopkins IRB when they let the study go on without crucial revisions in its design.

In the next section, we make the case that the problem was not a case of anomalous IRB error, but rather a function of the limitations of the institutional structure and makeup of IRBs.

### **The limitations of IRBs and institutional science ethics**

“Do no harm.” This most fundamental and general of ethical principles was not enacted by either the KKI researchers or compelled by the IRB reviewing their work. Why? It is evident from the transcripts of the court case that the recruitment of families with small children to live in abated homes was part of the protocol submitted by the researchers to the IRB<sup>12</sup>, and not an afterthought by an overzealous research assistant.

Indeed, in this as in other IRB processes, vetting specifically included careful attention to how research subjects will be recruited.<sup>13</sup> The approval of the study thus rested on the questionable ethical assumption that the recruited kids would have lived in lead paint filled homes anyways (given the preponderance of lead problems in the area), and that at least now there would be some level of abatement and consistent monitoring.

This type of reasoning is based in standard criteria for science ethics. One of the traditional criteria for IRB review, one emanating from the *Belmont Report* and from federal legislation, is the risk-benefit ratio entailed in research. As demonstrated by the judgment of the Maryland Court of Appeals, the generally sensible requirement to weigh risk against benefit can easily lead to conflict among persons and groups with very different calculus. Moreover, it seems likely that persons who are weighing risk-benefit accruing to others (as an IRB must do) might not use exactly the same calculus as they would if weighing risk-benefit to themselves or to their family. We return later to this point.

Could the Johns Hopkins IRB have done a better job, perhaps preventing the case from going to society's ultimate arbiter of ethics, the court system? Perhaps not. Undoubtedly most IRBs have members capable of sensitive ethical judgments. But when the ethical boundaries are gray rather than black and white, when decisions require complex uncertain assessments of risk-benefit and, perhaps most important, when the people having to deal with the consequences of those judgments have very different life experiences than the people making them, the legitimacy of the judgments of even the most ethically sensitive IRB member is subject to question. There is just no substitute for *identification and empathy* with the people who will be exposed to the research, just as there is no substitution for engaging in communication with the groups from which subjects are to be recruited. In order to fully develop this claim, and to make the case that "new and improved IRBs" with more highly trained personnel or more or better rules will not suffice, we draw on theories of bureaucracy and bureaucratic procedure.

#### *The limits of rules and formalization*

In a recent issue of *Critical Care*, bioethicist Randi Shaul (2002) reviews the problems of research ethics committees against a history of unsatisfactory protections for research subjects:

*Vague notions of accountability, although well meaning and sentimental, offer little guidance to those they are intended to direct and little comfort to those they are intended to protect. . . [It] is imperative that the clarification of the roles and responsibilities of [research ethics committees] is not delayed.*

Shaul goes on to suggest a tightening of the rules for IRBs. In our judgment, Shaul provides a partially correct diagnosis and an incorrect, unworkable solution. Accountability is certainly one problem with all familiar approaches to institutionalized science ethics. But more rules will not provide a solution. In most instances of organizational failure, and especially when there is a perceived need

for greater accountability, the first response is to develop more rules, more precise rules, more standard rules, rules with greater reach, and more training in the rules. Often this does more harm than good, undermining existing rules, setting up conflicts among rules and making rule implementation more and more complex.

In decision-making about scientific research, as in virtually all forms of decision-making, a crucial question is the degree of formality desirable for effective decisions (Gross, 1953; Organ and Green, 1981; Schulz, 1998). Should decisions rely chiefly on detailed rules, i.e. programmed decision-making (Simon, 1957), or applied decision heuristics, or spontaneous response? At one extreme are decision processes that are entirely rule-based, covering every contingency, such that decisions are essentially self-executing. At the other extreme are decision processes that are barely processes at all in which decisions are made spontaneously based on immediate stimuli and response, relying on either habit or informal and perhaps even subconscious decision heuristics. Almost all group or organizational decisions can be found within the range of these two extremes. In most instances, a key to decision effectiveness is finding the optimal point between purely spontaneous decision-making and purely rule-based decision-making (Thompson, 1967). Recent research on red tape and bureaucratic procedures has given us much more insight into the specific mechanisms by which rules fail or succeed (e.g. Zhou, 1993; Bozeman, 2000), a trend that has continued unabated (e.g. DeHart-Davis and Pandey, 2005; Moon and Bretschneider, 2002; Scott and Pandey, 2000; Pandey and Scott, 2002; Bozeman and DeHart-Davis, 1999).

Many of these limitations of rules, including, instability of implementation and the changing organizational ecology of rules (see, Bozeman, 2000) seem relevant to IRBs and protection of human subjects. Indeed, research on IRBs has demonstrated instability in the implementation of rules (Goldman and Katz, 1982), especially as a function of the particular composition of committees (Lane, 2005) and some propensity for subversion of rules (Barber et al., 1973).

Even the simplest rules are subject to great variance in interpretation, enactment, implementation, and effect. The greater the complexity of the social phenomenon, all else being equal, the greater the fallibility of the rules. When rules deal with simple procedures (e.g. destroy records) they have some chance of proving efficacious. When rules deal with ideals their utility diminishes sharply (for an elaboration and case study of rule efficacy see Bozeman and DeHart-Davis, 1999). The KKI study illustrates this very clearly. "Informed consent" is in many ways an ideal in that we believe that when people have all of the information they will by and large make good choices. One of the problems is immediately evident: how will we know when potential subjects have "all of the information?" Typically IRBs do not have any means for answering that question (other than their "best guess" based on experience), and there is little evidence for any formal efforts by IRBs to ascertain that subjects do in fact have enough information.

More rules, better rules, more precise rules, more conditions for rules, does not adequately deal with the inherent subjectivity in interpreting rules that verge on ideals. Instead, we argue for a second look at just who is doing the interpreting, and on whose behalf.

*Are there alternatives to bureaucratic control?*

Rules are means of exerting control- they require or prohibit behavior. Sometimes bureaucratic control is an important goal and rules provide an excellent means of enhancing goal achievement. Indeed, it certainly seems to be the case that the rules developed in the National Research Act and the more specific rules developed by IRBs and federal agencies are, on balance, quite salutary. But we should not expect that new rules will prevent or even significantly reduce controversies akin to the KKI lead paint study.

Instead of more or improved rules we argue for improved organizational processes and, in particular, an expanded role for participants. However, our argument is less one of general public participation than greater participation by potential research subjects.

### **Improving institutionalized science ethics**

The IRB has a near institutional monopoly on certifying the ethics of research. Nevertheless, if we consider the spectrum of institutional and non-institutional means for protecting research subjects, it is clear that the IRB is not – and should not be considered – the sole guarantor of ethical research. The first line of defense against exploitative research is constituted by the ethics of individual researchers; the last line of defense, one that only occurs after harm has been inflicted, is appeal to the courts. The question we ask here is: “Are there institutional designs that will provide advantages that cannot be obtained within the framework of IRBs, individual researchers, and the legal system?”

In any attempt to improve decisions outcomes, three critical questions emerge: (1) what constitutes decision quality? (2) What are the decision rules and processes? (3) Who should participate in decisions? Currently, IRBs have a satisfactory working definition of a quality decision. A good outcome from an IRB decision is that the research that is undertaken protects vulnerable populations, provides confidentiality, includes informed consent and has a suitable risk-reward ratio. With respect to the second decision-making question, rules and processes, the limited evidence available (e.g. Lane, 2005) suggests great variety in IRB processes and some considerable diversity in the use and implementation of rules. If there are shortcoming in rules and processes, these are probably issues of IRB reform rather than a call for new institutional designs.

It is the third decision-making question that suggests the need for an alternative science ethics institution. The answer to the “who participates” question for IRBs is that “it depends, but the participants will be few and they will not adequately represent the research population.” The structure of the IRB ensures that the social and demographic attributes of IRB members will closely resemble the researchers’ and will not much resemble the research subjects, especially when those subjects are deemed a vulnerable population. Realistically, it cannot be otherwise. To function expeditiously there must be some strict limits to the size of the IRB and there must be sufficient common ground and technical expertise to make informed decisions.

The fact that IRB committees do not bear a great resemblance to certain vulnerable populations that are potential research subjects is partly mitigated by two factors. First, the process of deliberation fostered by the IRB, regardless of its makeup, is likely to bring to the surface more considerations than would be dealt with by an individual. Second, IRB committees do strive for and often achieve a modicum of diversity in their make-up, whether of race, gender, or profession. Nevertheless, it is important to note that poor people – the subjects in the KKI study and in many other studies – are universally underrepresented on IRBs. While expanding the ranks of minorities in science in general and on IRBs in particular is an important and laudable goal, it will not bridge cultural and class divides IRBs are charged to make judgments across.

What this entails, then, is that the participative element of decisions cannot be easily accommodated by IRB reform; rather an altogether different institution is required.

### *Science ethics institutions: More participation as an alternative to more rules*

“Participation in science” often serves as mantra and its adherents sometimes do not bother to provide an explanation as to why participation in science is useful or in what forms it may be useful. There are, indeed, many important reasons to value participation in science and others have explored participation issues in detail (e.g. Schensul, 2002; Zaal and Leydesdorff, 1987; Chopyak and Levesque, 2001). Our concern is limited: why would broader participation improve decisions about the ethics of particular research projects and proposals?

In answering this question, we return to the KKI case. In the words of one of the attorneys, their plaintiff, “signed the informed consent, but no one ever told her, ‘there’s lead in this house, and it can cause brain damage,’” (Lewin, 2001). The researchers and KKI vehemently defended both the means and ends of the research. The chief executive of Kennedy Krieger, Gary Goldstein, maintained, “The impression of everyone doing the study was that everyone understood the situation,” (Lewin, 2001) and that participants in the study were “provided a package of benefits... they would not have had otherwise” (Rodgers and Stephenson, 2001).

Evidence suggests that (1) the research subjects *did not* have this information, at least not in a comprehensible form, and (2) the IRB *did* have the information. In our judgment, the wrong people had the right information. This seems to echo the Court’s opinion concerning this issue. In addition to denouncing the implementation of informed consent in the KKI study, the Court mandated that parents in the state of Maryland could no longer give consent for their children to participate in non-therapeutic research that poses any degree of risk to the children (Ross, 2002). In our view, the Court decision sometimes reacted too strongly to accepted research practice. But just considering subject recruitment and provision of informed consent, *does anyone doubt that if the KKI researchers had presented in full their research protocol to the research subjects, in a manner than was fully understood by the research subjects, that they would have said, “sure, sign us up to be exposed to lead poisoning?”* The problem was that there was no “jury of peers,” people who could easily see themselves in the same circumstance, to say “hold on!” The KKI informed consent may have passed the IRB test but it seems highly unlikely that it would have passed the scrutiny of a “vulnerable population,” had that population understood

fully the risks involved and, just as important, had that set of decision-makers been potentially part of the vulnerable population.

What is missing from IRB procedures, and cannot be included without fundamentally changing the institution, is empathy. Among the many possible reasons for enhancing participation in science, the need for direct representation is the greatest one (Laird, 1993) especially when one considers, in addition to empathy and representativeness, implications for civil society (DeLeon, 2002).

*Participant review boards: A participative approach to science ethics*

Our proposal is simple: let the IRBs do what they do well and consider other institutional approaches to attacking those problems not easily addressed by IRBs or similar bureaucracies. What IRBs do well is the “insider work,” the kinds of work that require some general passing knowledge of scientific work, experimentation and even research design. The rule for the use of professional communities, peers, and IRBs should be the same: they should work on the technical elements of ethical problems, deciding, in the above example, not about the need for confidentiality, but about the specific procedures for its protection. Specialists have no more appreciation for confidentiality than anyone else does and no better ability to predict harmful effects once it is breached. However, specialists do have a greater understanding of researchers’ behaviors with respect to data, the forms in which data are stored, and the statistical ability to infer from data. To sum up, IRBs are rule-bound, technical organizations and they should do what such organizations can do well, little of which has to do with either empathy or representation.

As a complement to “Institutional Review,” we propose the establishment of “Participant Review Boards” which could become involved only in those IRB cases requiring highly consequential decisions about the ethics of research and, particularly, the impact of research on vulnerable populations. *The basic idea behind a Participant Review Board is that the decision making process would involve persons resembling in every respect feasible the intended subjects of the research and who thus can identify, represent and communicate issues of concern for the subject population.* Our approach resembles the suggestions of Surowiecki’s (2004) *Wisdom of Crowds*, which argues that very large groups of “ordinary” people are likely to make better decisions and better forecasts than small groups of experts. We do not feel that very large groups of people necessarily make better forecasts, at least not within the domain of scientific research, but we do feel that they make “better” (more legitimate and more representative) ethical decisions, especially when they are making decisions that could be directly consequential for themselves. Our case for relying on the large, representative groups is not that they make more accurate or precise forecasts but that their ethical decision are better *by virtue of being representative*.

The Participant Review Board (PRB) approach, which shares some of the features of opportunities for public comment that are in use by agencies such as the EPA, should be used in cases where the stakes are large and the potential for both benefit and harm are substantial. In more routine cases, IRB procedures are acceptable and sufficient, but in cases where the relative weighting of benefits and costs depends on subjective interpretation across class and cultural divides, Participant Review Boards can fill the

gap. The logic, whereby an IRB faced with a protocol that involves a tricky balancing of costs and benefits to a vulnerable population can invoke the engagement of another party, is consistent with Code of Federal Regulations, Title 45 CFR Part 46.407.<sup>14</sup> This portion of the Common Rule allows IRBs to refer research that purports to help children but may pose risks greater than normally acceptable to the department of Health and Human Services for further review.

#### *Participant review board procedures*

We present these procedures as illustrative. We are proposing an approach that has not, as far as we have been able to determine, been used, at least not in connection with scientific research and human subjects. The procedures:

1. Before a study is authorized, present a fully-developed, fully-communicated research protocol to a panel composed of: (1) a designated advocate for the study; (2) a designated opponent of the study; (3) a designated mediator (by definition seen by all participants as neutral); and (4) a “PRB” that is, essentially, an acceptable ‘comparison group’ group for the proposed study- that is, to the extent possible, they are like the intended experimental subjects in every way except that they will not be participants in the study.<sup>15</sup>
2. The comparison group is instructed that the overarching criterion is “would I participate in this study or would I permit a loved-one to participate in this study?”
3. The mediator is responsible for allowing the “process” to take place. To do this he or she would provide opportunities for both sides (the advocate and the opponent) to make their cases. The mediator would be charged with questioning both sides to clarify their points, and most important, the mediator would enable the comparison group to question advocates, proponents and witnesses, and would be responsible for insuring that both the questions and answers would be understood by all involved. In effect, the mediator is a “specialist” in facilitating communication between diverse groups on complex issues. It is this process that is likely to be absent in the IRB process.
4. The comparison group would deliberate independently, but with the ability to review evidence and submit further questions.
5. The final step is for the comparison group to decide whether the research is permissible as is, whether it would be permissible with modification, or whether it is in some respect objectionable and, thus, not permissible.

These procedures are extraordinary efforts to be employed with extraordinary projects-defined as having a known and significant potential for harm, but which are believed to be justified by a known and substantial potential for benefit. Few studies fit in this category. Moreover, the *spirit* of the Participant Review Board approach is generalizable and includes the following: (1) A recognition that scientists have *less* authority to make value judgments about harm and benefit than does a group of people having attributes identical to or similar to the intended human subjects; (2) A recognition that the most important decisions about research ethics will be

improved by deliberation and, in some cases, by dialectic; (3) A recognition that standards and judgments are fluid and changeable; (4) A recognition of the intrinsic authority of communities, at least when persons from a recognizable community are intended to either suffer a differential burden or sustain a differential benefit; (5) A recognition that effective communication between the scientists and their subjects is vital in the process of deciding whether or not a study should go forward; and (6) A recognition that there are situations in which final decisions about research projects lie with members of the subject population and not with the scientific community.

*Science ethics institutions from a decision perspective*

The value of the PRB institution lies in large measure in the fact that it differs so much from the IRB in basic assumptions, procedures and type and source of legitimacy. As a decision-making institution it plays a role that the IRB cannot. To make this point, consider Figure 1.

Figure 1 depicts science ethics institutions on two axes, one representing the breadth of participation in decisions and the other the extent to which decisions are rule-bound (not the extent of rules but the extent to which decisions are “programmed” by rules). Science ethics decisions were, in the 1950’s, largely within the purview of the individual investigator, i.e. minimal participation, and the decisions may have included great ethical deliberation but not ones that were rule bound (at least not externally imposed rules). The situation has changed for the PI today, with decision-making being less unitary and more rule bound. Moreover, the PI and her ethical deliberations are formally linked to those of the IRB, a rule bound organization that includes more participants.

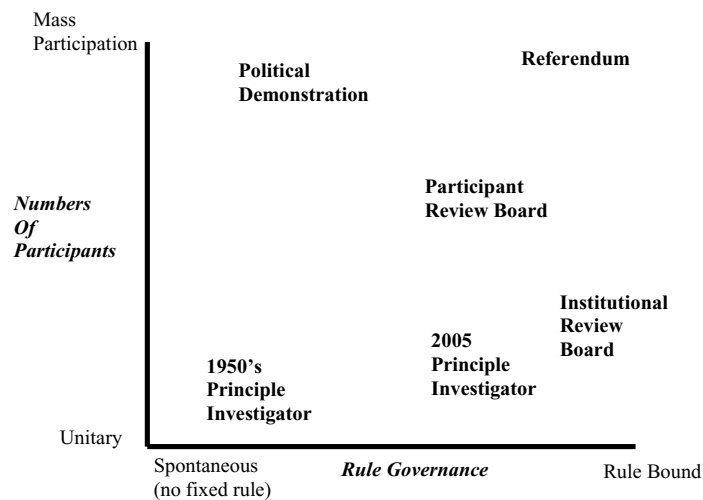


Fig. 1. A decision perspective on science ethics institutions.

The PRB approach provides for much broader participation, including many from the subject population or its representatives, as well as scientists and advocates. It is not so rule bound as either an IRB or a referendum, but does play by known rules. Moreover, it brings a different sort of science ethics decision-making entity to the table, one that receives its legitimacy from representation and proceeds on the basis of a structured dialectic. IRBs approach science ethics starting with ideals (e.g. protecting confidentiality, informed consent) and apply cases to those ideals. The procedure in the PRB system is more pragmatic, ultimately relying on the use of information for persuasion and the use of representation and the vote for its legitimacy. The PRB and the IRB each proceed on the basis of important values and procedures, but quite different and potentially complimentary ones.

#### *Limits of participant review boards*

It is possible to anticipate some of the objections to our PRB approach. One of the most obvious drawbacks is cost. The cost of a “fully burdened” IRB, one that includes the time for preparation of forms, the time of committee members, and indirect costs, is substantial but certainly much less than the cost of a PRB. The cost is not a good argument against the use of PRBs but a good argument that they should be used sparingly. Moreover, the costs are almost certainly less than trials in civil courts, even if one does not include damages awarded. Our suggestion for keeping down the cost of ethics institutions, which we do not elaborate here, is to greatly reduce the number of studies taken to IRB reviews. As we mentioned above, when not dealing with vulnerable populations, an informed consent form should suffice in a great many cases. It is beyond our scope to document the extraneous use of IRB review and the attendant compliance costs, but this would be a useful research agenda.

A greater difficulty of the PRB proposal lies in determining when to use the extraordinary measures of a PRB. One approach is that researchers undertaking controversial work could request a PRB. Government agencies could make a PRB a requirement for certain types of work. A more routine approach might be to conduct pilot research. It should be possible, especially in light of the expertise resident at universities, to devise opinion research items gauging the reaction of vulnerable populations to research that is about to be conducted on them. If a preliminary survey shows considerable divergence of opinions or obvious misunderstanding, then there is good prima facie evidence for the convening of a PRB.

Perhaps most importantly, we must understand the implications of substituting the judgments of well educated persons, many of whom are trained in research, with the judgments of persons who in some instances will be poorly educated, economically disadvantaged, and may have no understanding of the assumptions underlying scientific research. One might ask: how could such a body of people make “good” decisions? The obvious reply is that if there were no way to adequately inform a comparison group so that they are deemed able to make sound decisions, wouldn't this mean that informed consent for the actual research would be by definition impossible? As we have discussed, greater scientific knowledge, more education, and even grounding in moral philosophy do not necessarily yield ethical decisions. A doctoral degree does not make anyone more ethical and it does not make one's values more

sensitive and refined. Our proposal is that an appropriate group of people, one from the same environment and potentially facing the same risks as the would-be subjects, can provide the most valid answer to the question “should X harm be permitted for Y benefit?”

Of course, as Surowieki and others point out, not all crowds are wise, and in order to maximize the “wisdom” of PRBs it is important that several conditions are met. For Surowieki, a key condition is that decision-makers are able to make judgments independent of the undue influence of others. While a participant review system may alleviate some aspects of the power imbalance between researcher and vulnerable subjects, they may simultaneously create other problems of power and influence. We must be aware, for example of the possibility that by giving members of the PRB “authority” to protect the interests of potential subjects we may be changing their perspective in subtle ways; it is not clear whether they would as a result be more or less cautious in deciding what is and is not potentially harmful to subjects.

If we are to undertake such an endeavor it will therefore be crucial to invest adequate resources so that people participating in a PRB are truly empowered to access their individual and collective wisdom about what will inevitably be some complicated and touchy issues. Among other things, this means applying effort to insure the strength of arguments made *against* carrying out proposed research. One of the advantages of the “mediation model” is that this process is designed so as to insure that imbalances in power are minimized and that each side is empowered to speak freely on its behalf. Despite its difficulty, we believe that this aspect of the process is essential, and is ultimately one of the strengths offered by our method. After all, implicit is the idea of informed consent is the possibility and validity of saying no.

A likely outcome of a PRB operating under the procedures we suggest is that some useful research that would meet with approval under current institutions would be vetoed. In all likelihood, the important work represented by the KKI lead paint case might not have been accomplished. We do not wish to make it more difficult for science to address the needs of the disadvantaged. However, the disadvantaged deserve a vote, more than a token one; if they are to serve as the subjects, they deserve the deciding vote. We do not feel that it is overly optimistic to say that the adoption of a PRB institution for making some decisions about research ethics could have extremely positive long run effects in promoting trust between scientists and their subjects and in reducing the cultural divide. A potential outcome is that subject populations would be more willing to take on risks if they are clear about the importance of the study to members of their own community, and not just to the scientific community, and if they genuinely believe that they have the opportunity to say “no.”

A further bonus to the PRB system: if scientists show up for the proceedings, ordinary people will learn something about what often seems a black art, and scientists will certainly learn something about ordinary people, outside the courtroom, that is. Finally, this process could inform scientists about the importance of learning how to talk about their work in such a way that others – including of course those in the populations we have been talking about – would understand both its strengths and its limitations.

## **Conclusion: More rules or more participation?**

Let us consider a “thought experiment.” If the KKI study had employed a PRB approach, what would have been the result? In the first place we can probably have some confidence that the KKI case would have been deemed as one of those that is sufficiently controversial, at least potentially, to warrant a full blown set of procedures, complete with representation via a comparison group. It would not have “slipped through the cracks.” Would a better decision have been the result? That is not at all clear because we cannot say that the IRB decision in the KKI case were “bad decisions.” If the criterion for effectiveness of the IRB decision is that the benefit-risk ratio was acceptable, then it seems plausible that this was a good decision. But if effectiveness criteria included maintaining trust between the research community and the human subject community, in this case a set of vulnerable human subjects, the decision was perhaps not a “good decision.” (Though even from this perspective it is still not evident what the majority of families participating in the research felt about KKI procedures and outcomes- what we know is from those who decided to become plaintiffs).

It seems likely that a fully-executed PRB procedure would have had one of two results. In the first place, the research may have been stopped, or seriously altered, before it started. Second, if the research were approved, it seems likely that “informed consent” would have been more valid and that the research, once legitimated by a community of potential or actual participants, would have less likely resulted in either court action, scathing public opinion, or increased mistrust. In short, the PRB procedures seem unlikely to greatly improve the scientific merit of studies and likely to prevent at least some studies that would otherwise be authorized. Would the result be some sort of Bentham-like optimum? Probably not. However, if one’s long-run interest is less in the value of particular studies, no matter how great their potential merit, than in the value of retaining some faith in the institutions of science and in the democratization of science, then perhaps a new set of participative approaches could improve upon the current IRB institutional monopoly. It seems to us quite unlikely that clearer rules, more specific rules, or even better and more consistent rule enforcement could confer benefits related to the increased legitimation of the research or increased trust between researchers and the “researched.”

## **Notes**

1. <http://www.hopkinsmedicine.org/press/2001/SEPTEMBER/leadfactsheet.htm>
2. Ibid.
3. Many aspects of the legal code may affect science ethics (e.g. product liability laws, criminal law), but we consider as part of “institutionalized science ethics” only those statutes, rules and regulations directly aimed at research and researchers.
4. The WMA was created at a meeting held at the British Medical Association in London in 1946, and its first meeting of its General Assembly was convened in Paris in 1947, just a few weeks after the conclusion of the aspects of the Nuremberg Trial dealing with criminal experimentation on human subjects (Fluss, 2000). We can assume that those drafting the Declaration were influenced by the trial and the Nuremberg Code and, indeed, some of its language is nearly identical. However, it was not until 1953 that the idea for a position paper on the topic of human experimentation was first discussed

by the WMA's Medical Ethics Committee. Despite the time gap between the Nuremberg Code and the WMA declaration, Rothman (1998, p. 51) observes that "Nuremberg was the critical precedent for the format and collective pronouncements of such bodies as the American Medical Association and the World Medical Association. What began outside the domain of medicine came, in time, to be integrated in medicine."

5. In 1932, 399 African American sharecroppers living in Macon County, Alabama were diagnosed with syphilis, but they were not informed of the diagnosis. Instead, they were offered free meals, medical exams, and burial insurance to take part in a research study, funded in large part by the US Public Health Service, on the effects of untreated syphilis on the African American male. Although the men had agreed to be examined and "treated," they were not informed of the real purpose of the study. Even after penicillin came to be considered an effective treatment for syphilis in the 1940s, treatment continued to be purposely withheld from the research subjects so as not to interfere with the disease's natural progression. Amazingly, the study continued without any significant changes in the methodology until 1972, after a series of news articles condemning the studies were published (Jones, 1993; CDC, 2004).
6. In the face of growing demand for institutionalized science ethics, private, fee-for-service IRBs have emerged. These private IRBs have created some controversy about possibilities for bias and conflict of interest (Lemmens and Freedman, 2000).
7. Interview with former head of the IRB committee at a major research University [Note: Because of the controversies associated with the KKI case, we do not identify by name any of our interviewees].
8. Ibid.
9. A lead researcher we interviewed for this paper, who is on the faculty at Johns Hopkins University, bemoaned the fact that so little progress has been made on the lead paint problem. In pointing out that one justification put forth by landlords and politicians for doing nothing is cost, the researcher strongly advocated the need for a *real* study to get at the crucial question: what is the cheapest treatment that makes houses safe to live in? The KKI study was intended as a first phase of such a "real study"
10. <http://www.hopkinsmedicine.org/press/2001/SEPTEMBER/010907.htm>, accessed August 2005.
11. Anonymous interview with lead paint researcher on the faculty of Johns Hopkins University
12. "To help insure that study dwellings are occupied by families with young children, City Homes [a non-profit entity that owned and managed inner-city Baltimore Homes] will give priority to families with young children when renting the vacant units following Repair and Maintenance interventions." (Grimes vs. Kennedy Krieger, p. 3).
13. Interview with former head of IRB committee at a major research University.
14. [http://www.hhs.gov/ohrp/children/guidance\\_407process.html](http://www.hhs.gov/ohrp/children/guidance_407process.html), accessed Nov. 2005.
15. In an ideal scenario, the comparison group would be selected from the same sampling frame as the intended research subjects.

## References

- Adelman, S. A. (2001). *Appellee's Motion for Partial Reconsideration and Modification of Opinion; Grimes vs. Kennedy Krieger Institute, Inc.; and Higgins vs. Kennedy Krieger Institute, Inc. September*.
- Agency for Toxic Substances and Disease Registry (ATSDR), 'US Department of Health and Human Services' (1988). *The Nature and Extent of Lead Poisoning in Children in the United States: A Report to Congress*. Washington, D.C.
- Barber, B., J. Lally, J. Makarushka and D. Sullivan (1973). *Research on Human Subjects: Problems of Social Control in Medical Experimentation*. New York: Russell Sage Foundation.
- Bell, J., J. Whiton, and S. Connelly (1998). *Final Report: Evaluation of NIH Implementation of Section 491 of the Public Health Service Act, Mandating a Program of Protection for Research Subjects*. Arlington, VA.
- Bernard, S. (2004). 'Should the centers for disease control and prevention's childhood lead poisoning intervention level be lowered,' *American Journal of Public Health* 93: 1253-1260.
- Bozeman, B. and L. DeHart-Davis (1999). 'Red tape and clean air: Title V air pollution permitting implementation as a test bed for theory development', *Journal of Public Administration Research and Theory* 9: 141-177.

- Bozeman, B. (2000). *Bureaucracy and Red Tape*. Upper Saddle River, New Jersey: Prentice-Hall.
- Brown, G. E. (1993). 'The mother of necessity: Technology policy and social equity', *Science and Public Policy* : 411–416.
- Campbell, G. (2000). 'United States Demographics', in G. C. Jr., R. Denes and C. Morrison, eds., *Access Denied: Race, Ethnicity, and the Scientific Enterprise*. New York: Oxford University Press, pp. 7–41.
- Center for Disease Control (2004a). *Timeline- The Tuskegee Syphilis Study: A Hard Lesson Learned*.
- Center for Disease Control (2004b). *General Lead Information: Questions and Answers*.
- Chopyak, J. and P. Levesque (2001). 'Public participation in science and technology: Decision-making: Trends for the future', *Technology in Society* 24: 155–166.
- Cleary, R. E. (1987). 'The impact of IRBs on political science research', *IRB* 9: 6–10.
- Curry, D. (2001). 'Court condemns Hopkin-approved study that exposed children to lead dust', *Chronicle of Higher Education* 7: A32.
- DeLeon, P. (2002). 'Policy analysis in the good society', *The Good Society* 11: 37–41.
- DeHart-Davis, L. and S. K. Pandey (2005). 'Red tape and public employees: Does perceived rule dysfunction alienate managers?,' *Journal of Public Administration Research and Theory* 15: 133–148.
- Dunn, C. and G. Chadwick (1999). *Protecting Study Volunteers in Research: A Manual for Investigative Sites*. Boston, MA: CenterWatch, Inc.
- Enserink, M. (2000). 'Helsinki's new clinical rules: Fewer placebos, more disclosure', *Science* 290: 418–419.
- Farmer, P. (1999). *Infections and Inequalities: The Modern Plagues*. Berkeley: University of California Press.
- Ferraro, F., E. Szigeti, K. Dawes, and S. Pan (1999). 'A survey regarding the University of North Dakota Institutional Review Board: Data, attitudes, and perceptions', *Journal of Psychology* 133: 272–280.
- Fluss, S. (2000). 'How the declaration of Helsinki developed', *Good Clinical Practice Journal* 6: 18–22.
- Goldman, J. and M. Katz (1982). 'Inconsistency and Institutional review boards', *The Journal of the American Medical Association* 248: 197–202.
- Gray, B. H. (1978). 'Institutional review boards as an instrument of assessment: Research involving human subjects in the U.S.,' *Science, Technology, and Human Values* 4: 34–46.
- Gray, B., R. Cooke and A. Tannenbaum (1978). 'Research involving human subjects', *Science* 201: 1094–1101.
- Gross, E. (1953). 'Some functional consequences of primary controls in formal work organizations', *American Sociological Review* 18: 368–373.
- Howard-Jones, N. (1982). 'Human experimentation in historical and ethical perspectives', *Social Science and Medicine* 16: 1429–1448.
- Human, D. and S. Fluss (2001). 'The world medical association's declaration of Helsinki: Historical and contemporary perspectives', <http://www.wma.net/e/ethicsunit/helsinki.htm>. Accessed July 29, 2004.
- Johnson, E. B., S. Jackson Lee, D. Hooley, and B. Lee (1998). *Dissenting Views to Unlocking the Future: Report of the House Science Committee Science Policy Study*. Washington, D.C.
- Jones, J. (1993). *Bad Blood: The Tuskegee Syphilis Experiment*. New York: Free Press.
- Kaiser, J. (2001). 'Human subjects: Court rebukes Hopkins for lead paint study', *Science* 293: 1567–1569.
- Laird, F. (1993). 'Participatory analysis, democracy, and technological decision making', *Science, Technology and Human Values* 18: 341–361.
- Lamphear, B. and M. Weitzman, et al. (1996). 'Racial differences in urban children's environmental exposure to lead', *American Journal of Public Health* 86: 1460–1463.
- Lane, E. (2005). *Decision-Making in the Human Subjects Review System*. Unpublished doctoral dissertation. Atlanta, GA: School of Public Policy, Georgia Institute of Technology.
- Lemmons, T. and B. Freedman (2000). 'Ethics review for sale? conflict of interest and commercial review boards', *Milbank Quarterly* 78: 547–562.
- Lewin, T. (2001). 'US investigating John Hopkins study of lead paint hazard', *The New York Times* August 24.
- Moon, M. J. and S. Bretschneider (2002). 'Does the perception of red tape constrain IT innovativeness in organizations? Unexpected results from a simultaneous equation model and implications', *Journal of Public Administration Research and Theory* 12: 273–291.

- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979). *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research*. Accessible at: <http://ohsr.od.nih.gov/guidelines/belmont.html>.
- National Research Council (1993). *Measuring Lead Exposure in Infants, Children, and Other Sensitive Populations*. Washington, D.C.: National Academy Press.
- National Institute of Health (2005) *Title 45 CFR Part 46 Protection of Human Subjects*, Office of Human Subjects Research, <http://www.nihtraining.com/ohsr/site/guidelines>. Accessed May 6, 2005.
- Organ, D. and C. Greene (1981). 'The effects of formalization on professional involvement: A compensatory process', *Administrative Science Quarterly* 26: 237–252.
- Pandey, S. K. and P. G. Scott. (2002). 'Red tape: A review and assessment of concepts and measures', *Journal of Public Administration Research and Theory* 12: 553–580.
- Rodgers, J. and G. Stephenson. (2001). *Kennedy Krieger, Hopkins, UMMS and Others Ask Court of Appeals to Reconsider Issues in Lead Paint Study Decision*. John Hopkins Medical Institute. Sept. 17, 2001.
- Roig-Franzia, M. and R. Weiss (2001). 'Md. appeals court slams researchers', *Washington Post*. August 20, 2001. Front page.
- Ross, L. F. (2002). 'In defense of the Hopkins lead abatement studies,' *Journal of Law, Medicine, and Ethics* 30: 50–57.
- Rothman, D. J. (1998). 'The nuremberg code in light of previous principles and practices in human experimentation', in U. Tröhler and S. Reiter-Theil, eds., *Ethics Codes in Medicine: Foundations and Achievements of Codification Since 1947*. Aldershot, England, and Brookfield, VT, USA: Ashgate Publishing, pp. 50–59.
- Scott, P. G. and S. K. Pandey (2000). 'The influence of red tape on bureaucratic behavior: An experimental simulation', *Journal of Policy Analysis and Management* 19: 615–633.
- Schulz, M. (1998). 'Limits to bureaucratic growth: The density dependence of organizational rule births', *Administrative Science Quarterly* December, 1998.
- Shaul, R. (2002). 'Reviewing the reviewers: The vague accountability of research Ethics committees', *Critical Care* 6: 121–122.
- Shensul, J. J. (2002). 'Democratizing science through social science research partnership', *Bulletin of Science, Technology, and Society* 22: 190–202.
- Simon, H. A. (1957). *Administrative Behavior*. New York: Mac Millan.
- Spriggs, M. (2004). 'Canaries in the mines: Children, risk, non-therapeutic research, and justice', *Journal of Medical Ethics* 30: 176–181.
- Surowiecki, J. (2004). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*. New York: Doubleday.
- Thompson, J. (1967). *Organizations in Action*. New York: McGraw-Hill.
- U.S. Food and Drug Administration (2005). *Guidelines for Institutional Review Boards and Clinical Investigators* <http://www.fda.gov/oc/ohrt/irbs/faqs.html#IRBOrg> Accessed May 10, 2005.
- Zaal, R. and L. Leydesdorff. (1987). 'Amsterdam science shop and its influence on university research: The effects of ten years of dealing with non-academic questions', *Science and Public Policy* 14: 310–316.
- Zhou, X. (1993). 'The dynamics of organizational rules', *American Journal of Sociology* 98: 1134–1166.